DOCUMENT RESUME

ED 354 067 PS 021 032

AUTHOR Halford, Graeme S.; Stewart, J. E. M.
TITLE Cognitive Science Questions for Cognitive Development: The Concepts of Learning, Analogy, and Capacity.
PUB DATE 92
NOTE 38p.; An earlier version of this paper was presented at the International Congress of Psychology (25th, Brussels, Belgium, July 19-24, 1992).
PUB TYPE Information Analyses (070) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Development; Cognitive Mapping; *Cognitive Processes; Developmental Stages; Foreign Countries; Individual Differences; *Learning; Learning Problems; Learning Strategies; *Preschool Children; Preschool Education; Schemata (Cognition); Serial Ordering
IDENTIFIERS Analogical Reasoning; *Cognitive Models; *Cognitive Sciences; Conceptual Frameworks; Knowledge Acquisition; Parallel Processing

ABSTRACT

New conceptions of learning, analogy, and capacity have fundamentally changed scientists' view of cognitive development. New conceptions of learning help to explain how representations of the world are acquired. New models of analogical reasoning have suggested that logical inferences are often made by mapping a problem into a mental model, or schema, induced from ordinary life experience. A model of analogical reasoning provides a basis for understanding children's limitations in cognitive capacity, and specifies changes in the nature of children's cognitive representations over time that explain phenomena previously attributed to developmental stages. The concepts children understand, and the strategies they develop based on their understanding, depend on the complexity of the representations they construct. Parallel Distributed Processing (PDP), a model of cognitive processing, explains why the number of dimensions, or independent items of information required to represent a concept, that can be processed in parallel is limited. The PDP model provides an account of the effect of the complexity of a concept on children's cognitive performance. In this model, cognitive growth depends on four main factors: (1) learning and induction; (2) conceptual chunking; (3) serial processing strategies; and (4) the development of the ability to represent concepts of higher dimensionality. A list of 39 references is attached. (MM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED354067

Cognitive Science Questions for Cognitive Development: The Concepts of Learning, Analogy, and Capacity

Graeme S. Halford and J.E.M. Stewart

University of Queensland, Australia

An earlier version of this paper was presented to the 25th International Congress of Psychology, Brussels, July 19-24, 1992. It is our pleasure to acknowledge the contribution made by numerous of our colleagues to the formulation and elaboration of this theory, and asociated work . They include John Bain, Campbell Dickson, Ross Gayler, Jian Guo, Bill Wilson, and Janet Wiles. Author's address: Graeme Halford, Psychology Department, University of Queensland, 4072, Australia. Fax: 61 7 365 4466. Email: gsh@psych.psy.uq.oz.au.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
G. S. Halford
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PS 021032

BEST COPY AVAILABLE

Abstract

New concepts from cognitive science have fundamentally changed our view of cognitive development. In this paper we explore the implications of three concepts from cognitive science. These are learning (and induction), analogy, and capacity. New conceptions of learning have enabled us to understand how representations of the world are acquired. New models of analogical reasoning have suggested that "logical" inferences are often made by mapping the problem into a mental model, or schema, induced from ordinary life experience. A model of analogical reasoning, based on neural nets, provides a natural basis for capacity limitations, and specifies changes in representations over age that explain phenomena previously thought to be stage-related.

Cognitive Science Questions for Cognitive Development: The Concepts of Learning, Analogy, and Capacity

The view of cognitive development that we wish to present can be summarized in the following propositions:

1. The concepts children understand, and the strategies they develop based on that understanding, depend on the complexity of the representations they can construct.

2. Conceptual complexity can be defined in terms of the number of independent dimensions that need to be represented. Parallel Distributed Processing models of the way information is represented help to explain why the number of dimensions that can be processed in parallel is limited. This leads to a new definition of processing capacity.

3. Some of the concepts that children find difficult require representations that exceed their processing capacity. This results in strategies that yield some correct solutions, but are not generally valid. This can account for many phenomena, including some that have traditionally been attributed to stages.

We would also like to conjecture that the phenomena which Piaget (1950) attributed to stages correspond to the number of vectors that can be processed in parallel in a parallel distributed representation (i.e. they are related to the rank of a tensor product of vectors). We will consider each of these points in more detail.

Representations have been defined in the cognition literature (Grossberg, 1980; Halford & Wilson, 1980; Holland, Holyoak, Nisbett & Thagard, 1986; Palmer, 1978), and their implications have been

summarized elsewhere (Halford, in press). The essence of a cognitive representation is that it consists of a cognitive structure which is in correspondence to a structure in the world. Structure is not used in the Piagetian sense, but means a set of elements on which one or more relations (or functions) is defined. Any aspect of the world can be thought of as a set of elements with relations between the elements. For example "human family" comprises the elements father, mother, child1, child2, ... These elements are linked by relations like "father of", "mother of", "sibling of", "sister of", "daughter of", and so on. A cognitive representation of family will comprise a set of internal mental elements, and a set of relations that correspond to those in the real world family.

It is not always necessary for the representation elements and relations to be the same as, or even to resemble, the real world elements and relations. It is sufficient for a representation to be valid if the structures correspond, for which mathematical definitions have been given (Halford, in press, Chapter 2; Holland et al., 1986). Therefore representations are not "pictures in the head".

Representations can take a number of different forms, but all comply with the criterion of structural correspondence. There has been considerable controversy about the reality of the distinction between images and propositional representations, but this issue need not concern us here because, as Palmer (1978) points out, two representations can be taken as equivalent if they contain the same information and (as Halford, in press), adds the information is equally accessible.

There is however one kind of representation that has important implications for both cognition generally and cognitive development in particular. This is the *parallel distributed processing* (PDP) approach to modelling the microstructure of cognition (Rumelhart & McClelland, 1986). According to this approach, representations consist of sets of units, each of which has an activation value. The set of activation values is normally expressed as a vector. There are excitatory and inhibitory links between the units, which effectively code the constraints, or regularities, operating on the structure. The links operate in parallel, so all constraints operate together. These representations have a number of important, if counter-intuitive, properties. These include;

1. Learning depends on changing the strengths of the links between units.

2. Representations can be superimposed on the same set of units.

3. The representations have emergent properties which include automatic generalization and discrimination, automatic averaging and prototype formation, and automatic regularity detection.

4. If units are lost the representation loses clarity but still functions to an extent that depends on the proportion of units remaining (graceful degradation).

5. If too many representations are superimposed on the same set of units, there is a loss of clarity resulting in ambiguity (graceful saturation).

6. There is no central control over processing, which consists in the representation "settling" into the state which best fits all constraints acting in parallel. Note that this means the distinction

between structure and process virtually disappears in some models of this type.

7. Recent proposals (e.g., Halford et al, in press: Hinton, 1990) have pointed the way to dealing with the integration of PDP representations into structures and concepts. Some of this work will be used in the proposals which follow.

Further summaries are given by Best (1992, chapter 7; Rumelhart & McClelland, 1986, chapter 1).

The importance of PDP models to our argument is that, as we will explain later, they lead to a natural account of the effect of conceptual complexity on performance. They also enable us to offer a new definition of processing capacity, and to redefine the question of whether capacity changes with age.

Conceptual complexity is not synonymous with difficulty. Tasks can be difficult for many reason besides their complexity. For example someone can fail a task for lack of knowledge or strategies (procedural knowledge), because of lack of availability of the correct hypothesis, poor motivation etc.

We define conceptual complexity in terms of dimensionality, which is the number of independent items of information required to represent the concept. Dimensionality is similar to the idea of degrees of freedom; i.e. the number of independent sources of variation in a particular system.

The general principles are:

1. Those variables that enter into the current computation must be represented, and;

2. Aspects of the situation which vary independently must be represented as separate dimensions.

The processing load for any step in a task corresponds to the number of dimensions that must be represented. It has been confirmed empirically that higher dimensionality is associated with higher processing load, with other factors controlled (Halford, Maybery & Bain, 1986; Halford & Leitch, 1989; Maybery, Bain & Halford, 1986).

As we will see, the number of dimensions can be linked to the number of vectors required to represent a concept in a PDP representation. This provides a natural explanation for the increase in processing load with concepts of higher dimensionality.

Processing capacity has proved a difficult and controversial topic, both in cognition and cognitive development. It has been considered in more detail elsewhere (Halford, in press, Chapter 3). There can be no doubt that cognitive development cannot be attributed solely to growth of capacity. There is too much evidence that many aspects of performance are attributable at least in part to accumulation or restructuring of knowledge (Carey, 1985; Chi & Ceci, 1987). However evidence of the importance of knowledge, skills, or strategies in no way denies that capacity may also play a role. Methodological difficulties have tended to prevent evidence

being obtained either for or against the proposition that capacity increases with age. However there is now a small but growing body of evidence that capacity does change with age. There is physiological evidence of capacity change (Diamond, 1989; Goldman-Rakic, 1987; Rudy, in press; Thatcher, Walker & Giudice, 1987). There is also evidence of a general processing speed factor that changes with age (Kail, 1991), and that primary memory capacity changes with age (Halford, Maybery & Bain, 1988).

Part of the problem is that the question has not been well defined, so researchers have not had a clear idea of what they were seeking. Capacity has often been identified with short term memory span, because of the memory theory of Atkinson and Shiffrin (1968) which implied that short term memory was the workspace of thinking. However as Baddeley (1990) has pointed out, there is little evidence to support this proposition, and there is considerable evidence that contradicts it. An extensive literature on working memory shows that there is little interference between various cognitive processes such as decision making or reasoning, and a concurrent short term memory task. See Baddeley (1990) or Halford (in press, Chapter 3) for reviews. If short term memory were the workspace of thinking, such interference would be expected. It seems more likely that short term memory depends on a specialized system, which Baddeley (1990) calls the phonological loop, and which is distinct from the central processor.

It appears that processing capacity should be distinct from storage capacity. Working memory is sometimes used to refer to information that is stored in short term memory for use in later

problem solving steps, but is not being currently processed. Ability to retain such information depends on storage capacity, but not on processing capacity. The latter term should be used for information that is currently entering into some kind of reasoning, decision making, or other computational process. Processing capacity is best defined in terms of the number of independent items of information, or dimensions, that enter into a specific computation.

Learning, and strategy development

If we accept that knowledge acquisition is a major component of cognitive development, it follows that learning, defined as acquisition of knowledge through experience, must play a significant role. Despite this there has been surprizingly little interest in the role of learning in cognitive development. Part of the reason is that the concept of learning is associated in the minds of psychologists with behavioristic learning theories which have not been found to offer any solutions to the problems of cognitive development. Such reservations, though very understandable in the past, are no longer justified, because there are contemporary learning theories which do have the potential to explain how children acquire important concepts, and are worthy of further study by cognitive developmentalists. The problem of learning has several aspects, each with associated theory, and We will consider each in turn.

The first aspect is acquisition of knowledge about the structure of the world. The cognitive representations discussed earlier comprise information about relations between things and events in the world, and this information has to be acquired. Given

the vast amount of information of this type that a child apparently acquires automatically, the learning process that is responsible must be very efficient. A reinterpretation of some established learning phenomena, including classical conditioning (Rescorla, 1988) and discrimination learning (Halford, in press, Chapter 4) shows that humans and (other) animals possess very basic and effective learning mechanisms for this purpose. Theories of this process have been proposed by Holland et al. (1986) and by Holyoak, Koh & Nisbett (1989). Furthermore PDP theory provides powerful explanations for our ability to extract regularities from experiences which include a lot of randomness.

The basic principles of this learning are that representations are strengthened when they validly predict relations between events, and weakened otherwise. Furthermore the strenghtening effect depends on the informativeness of the representation. Representations which make redundant predictions are not learned. These learning processes can go a long way towards explaining how children build up a store of knowledge about the structures, or relationships, in the world. These learned representations provide the "raw material" for the mental models that are increasingly being recognized as the basis of natural, human reasoning. These theories can do a lot to explain how knowledge becomes reorganized to meet the requirement for children to deal with increasingly complex and sophisticated concepts.

The second aspect of learning is acquisition of skills and strategies. There are sophisticated computational models of skill acquisition (Anderson, 1987) which can be applied to showing how

children acquire reasoning strategies. One such model (Halford, Maybery, Smith, Bain, Dickson, Kelly & Stewart, 1992; Halford, Smith, Dickson, Maybery, Kelly, Bain & Stewart, on contract) shows how transitive inference strategies can be acquired, and will be considered later. These models recognize the active, constructive role of the child in building its own knowledge base, and are a far cry from the passive associationistic theories of the past.

Transitive inference

We will explicate the theory of cognitive development through the task of transitive inference, which has been important throughout the history of cognitive development research, and for which a large, high quality data base has been assembled. Consider a transitive inference task such as: "Peter is fairer than Tom; John is fairer than Peter. Who is fairest (darkest)?"

There is a reasonable concensus in the literature that such tasks are performed by arranging the terms in order (Sternberg, 1980; Trabasso, 1977; Thayer & Collyer, 1978). However, before children's performance on this task can be understood, we need a conceptualization of the reasoning process.

It is becoming increasingly apparent that human reasoning is essentially analogical in character, particularly with novel problems. Therefore we can conceptualize transitive inference as mapping the premises into a schema, which is used as an analog, as shown in Figure 1.

Insert Figure 1 about here

In this case a common ordering schema, the top-down (or left-right) arrangement, is used as an analog. In effect it serves as a kind of template, or "mental model" for imposing order on the premises. Once the premises are ordered in this way transitive inferences are easily made by accessing the ordered representation; e.g. we can easily see that John is fairer than Tom.

There is some difficulty in performing the mapping however. This is because both premises must be processed to map any premise term into a slot in the ordering schema; e.g. we need both premises to know that John must go in first position. This illustrates a point which is of some importance to the theory, which is that mapping into analogs or mental models imposes a processing load, the magnitude of which depends on the complexity of the structure involved.

This analogical mapping process is important where a task is novel. Familiar tasks are usually performed using strategies acquired through past experience. Transitive inference strategies normally entail storing the premise terms as an ordered set in short term memory (Foos, Smith, Sabol & Mynatt, 1976). However development of these strategies depends on a concept of the task. According to our model, the concept of the task is based on a specific instance of an ordered set of at least three elements (Halford, et al., 1992; on contract).

We have developed a self-modifying production system model which acquires strategies through experience, guided by a specific example of an ordered set which is used as an analog, as shown in the previous transparency. Once such a strategy is developed there is no further need for analogical reasoning, except where the strategy must be modified, or transferred to a new domain.

One major goal of theory in this domain is to account for the difficulty which transitive inference tasks cause for young children. Attempts have been made to explain these difficulties away, on the grounds that they depend on flawed tests, producing false negatives, or lack of experience, resulting in inadequate knowledge.

However, many of the claims that children succeed with alternative tests are flawed due to either false positives (e.g. reporting chance results as success), or failure to consider alternative bases for the performance (Halford, 1989). Furthermore, many of the improvements have been with children *over* five years, and therfore do not account for the finding that these tasks are specially difficult for children below this age. Another problem is that lack of process models makes it difficult to define test validity, resulting in circularity; "good" tests tend to be those that children pass. Therefore it seems appropriate to conclude that while lot of important causes of failure have been discovered, but there are still sources of difficulty for young children that remain to be explained.

Therefore we must seek alternative explanations for the difficulties which children experience with transitivity and some

other tasks. Using the easy-to-hard paradigm of Hunt and Lansman (1982) we have found evidence that performance of children is capacity-limited on these tasks (Halford et al., 1986).

PDP implications for processing capacity

Some new insights into the basis of capacity limitations has been obtained from our work on parallel distributed processing models of analogies (Halford, Wilson, Guo, Gayler, Wiles, & Stewart, in press). This caused us to address the way concepts are represented in PDP architectures, and the approach we adopted leads to some insights into the reason why certain concepts are associated with high processing loads. We can examine this issue by seeing how concepts of varying complexities are represented. The representation of a binary relation, such as LARGER THAN is shown in Figure 2.

Insert Figure 2 about here

A vector is used to represent the predicate, LARGER THAN, and another vector is used to represent each argument. In this example, there is a vector representing arguments *elephant* and *dog*. The predicate-argument binding, that is, the fact that elephant is larger than a dog, is represented by the tensor product of the three vectors, as shown in Figure 2. Actually, each of the units in the vectors representing "larger than", "elephant", and "dog" is connected to one of the tensor product units in the interior of the figure, but the connections are not shown because they would make the figure too cluttered. The activations on these units effectively code the

relation between the vectors. The structure permits information about the relation to be recovered. Given the predicate and an argument we find possible cases of the second argument; e.g. given the predicate "larger-than" and "elephant" the representation permits retrieval of things (such as dogs) that are smaller than elephants, equivalent to asking what is smaller than an elephant? Alternatively, given the arguments, the predicate can be found, equivalent to asking what is the relation between elephant and dog.

Because LARGER-THAN is a binary relation, with two arguments, it is represented by a rank 3 tensor product, that is, one with three vectors. However more complex concepts are represented by structures with more vectors. The representation of transitivity requires a rank 4 tensor product, as shown in Figure 3.

Insert Figure 3 about here

Given that transitive inferences are made by organizing premise information into an ordered set of three elements, as shown in Figure 1, the core of the transitivity concept is a ternary relation. That is, transitivity is a relation with three arguments, corresponding to a,b,c or top, middle, bottom, depending on the particular instantiation.

Consequently, it has to be represented by a tensor product of higher rank than a binary relation, such as LARGER-THAN. A tensor product of higher rank imposes a higher processing load, because the number of tensor product units increases exponentially with the number of vectors, and the number of connections increases

accordingly. The PDP model therefore provides a natural basis for the increase in processing load that is observed with more complex concepts such as transitivity.

The rank of a tensor product can be shown to relate to the conceptual complexity metric based on dimensionality, as discussed earlier. Recall that the complexity of a concept is defined in terms of the number of independent items of information required for the computations the concept entails. The number of vectors required for a representation based on tensor products, according to the model of Halford et al. (in press) is one more than the number of dimensions. Hence a binary relation, which is 2 dimensional, is represented by a tensor product of rank 3. Transitivity is three dimensional, and is represented by a tensor product of rank 4. One advantage of the approach is that the rank of tensor product required for particular computations, and hence the dimensionality, can be confirmed by simulation.

Age and dimensionality of representations

This argument enables us to reformulate the longstanding question of whether processing capacity changes with age. The question becomes, not whether overall capacity changes with age, but whether representations become more differentiated so that tensor products of higher rank can be processed. This would mean that concepts of higher dimensionality would be represented, enabling higher-order relations to be understood. Representations of

varying dimensionality, with corresponding tensor products, are shown in Figure 4.

Insert Figure 4 about here

At the lowest level unary relations are represented. These are 1-dimensional concepts, and require tensor products of rank 2. They include simple categories, defined by one attribute such as the category of large things, or the category of triangles. They also include categories defined by a collection of attributes that can be represented as a single chunk, such as the category of dogs. One vector (shown vertically in Figure 4) would represent the category label DOG. The other vector would represent the instances. Representations of different dogs would be superimposed on this set of units. Thus vectors representing each known dog would be superimposed, so the resulting vector would represent the central tendency of the person's experience of dogs. It would represent the person's prototype dog. However the representations of the individual dogs can still be recovered. Questions such as "are chihuahuas dogs", or "tell me the dogs you know" can be answered by accessing the representation. Note that the representation is one dimensional because if one component is known, the other is determined. Thus if the argument vector represents a labrador, the other vector must be "dog". Similarly, if the predicate vector represents "dog", the argument vector must represent one or more dogs.

They also include ability to represent variable-constant bindings. The well-known A not-B error in infant object constancy

research can be thought of as requiring ability to treating hiding place as a variable. That is, when an infant has repeatedly retrieved an object from hiding place A, then continues to search for it at A despite having just seen it hidden at B, the infant is treating the hiding place as a constant. However if hiding place were represented as a variable this perseveration would be overcome. Thus the fact that the A not-B error disappears about one year is consistent with ability to represent rank-2 tensor products developing at that time. This implies that ability to construct representations equivalent to rank 2 tensor products probably develops at approximately one year of age. We would therefore predict that other performances which require this level of representation, should first appear at this time. There should be a general ability to represent variables as distinct from constants.

As we have seen simple categories also occur at approximately this age, and are represented by rank 2 tensor products. In general, the appearance of cognitions which require to be represented by rank 2 tensor products amounts to ability to relate one representation to another. The observations which Piaget attributed to the preconceptual stage appear to require this level of representation.

At the next level binary relations, and univariate functions can be represented. These are all 2 dimensional concepts (given any two components, the third is determined), and they entail tensor products of rank 3. Based on an assessment of the cognitive development literature Halford (1982, in press) suggests they develop at approximately two years of age. They correspond to

Piaget's observation that in the intuitive stage children process one binary relation at a time.

At the next level concepts based on ternary relations, binary operations, and bivariate functions, are represented. These are 3-dimensional, and require tensor products of rank 4. Well known examples include transitivity and class inclusion, but there are many other concepts that belong to this level, including conditional discrimination, the transverse pattern task, the negative pattern task, dimension checking in blank trials task, and many more (Halford, in press). The familiar binary operations of addition and subtraction belong to this level. One vector represents the operation (+ or x) while two others represent the addends (multiplicands), and the fourth vector represents the sum (product). Note that if you know three of these, the fourth is determined; e.g. if you know the numbers are 2,3,5 you know the operation is addition; if you know the numbers 2, ?, 5, and the operation is addition, you know the missing number is 3, and so on. (Readers interested in PDP might note that there is no catastrophic forgetting when addition and multiplication are superimposed on a rank 4 tensor product).

All of these tasks are performed by about five years of age, but cause considerable difficulty below this age. In a broad sense, this level of processing corresponds to Piaget's concrete operational stage, which can be conceptualized as ability to process binary operations, or compositions of binary relations (Halford, 1982, in press; Sheppard, 1978).

At the fourth level concepts based on quaternary relations, and compositions of binary operations, can be represented. These include understanding proportion and ability to reason about relations between fractions, as well as understanding concepts such as distributivity, that are based on compositions of binary opersations. In a broad sense this level of processing corresponds to Piaget's formal operations stage, which entails relations between binary operations (Halford, in press).

Chunks and dimensions

We have argued (Halford in press; Halford et al., in press) that the number of dimensions can be identified with the number of chunks. Miller's (1956) concept of a chunk is a unit of information that can vary in size. For example a letter, digit, or word, can all be chunks, even though they vary considerably in the amount of information they contain. The limit is in the number of chunks, irrespective of the amount of information. This entails a paradox, because the number of items is limited, but the amount of information is not. It means that the limitation is in the number of independent items that can be processed. One way to handle this is to compare chunks with dimensions. That is, a chunk, like a dimension, is an independent unit of information of varying size. It appears reasonable to identify chunks, and dimensions, with vectors,because each vector can represent varying amounts of information. Thus the explanation for the paradox may be that information is represented in vectors, each of which represents one chunk or one dimension.

Working memory research suggests that the number of chunks that adults process in parallel is about four (Schneider & Detweiler, 1987; Halford et al., in press). Therefore we would predict that adults can process a maximum of four dimensions in parallel. This would mean that the most complex tensor product representations that can be processed would be rank 5, i.e. with five vectors.

Chunking and segmentation

Concepts more complex than four dimensions can be processed by either *conceptual chunking* or *segmentation*. Conceptual chunking entails recoding concepts of higher dimensionality into fewer dimensions, most commonly into one dimension; i.e. it entails reducing multiple chunks to a single chunk. An example would be the concept of velocity, defined as v = s/t. It is 3 dimensional, and requires a tensor product of rank 4. However it is also possible to think of velocity as a single dimension, such as the position of a pointer on a dial.

When velocity is chunked as a single dimension, it can be represented by a single vector, and combined with up to three other dimensions. Thus velocity can now be used to define acceleration, a = (v2 - v1)/t.

Acceleration in turn can be chunked, and combined with up to three other dimensions. Thus force, F = ma can be defined as the product of mass and acceleration. Conceptual chunking enables us to bootstrap our way up to concepts of higher and higher

dimensionality, without exceeding the number of dimensions that can be processed in parallel.

If the number of dimensions can be reduced by chunking, is the limit in processing capacity meaningful? It is meaningful because when representations are chunked, we lose the ability to recognize relations within the representation. When velocity is represented as a single dimension, we can no longer compute the way velocity changes as a function of time or distance, or both. Similarly, we cannot compute what happens to time if distance is held constant, and velocity varies, and so on. This example illustrates the point that any computation requires a minimum number of dimensions to be represented.

Segmentation entails developing serial processing strategies. In this case tasks are segmented into steps, each of which is small enough not to exceed the capacity to process information. Only that part of a concept that is the focus of attention is represented at any one time. We are developing a model of this process in context of complex analogical reasoning. Complex analogies, such as that between heat-flow and water-flow, are represented by a hierarchical structure, in which an overall concept, such as that temperature difference causes heat-flow, is represented as a binary relation, without detail. At the next level down, the details of temperature difference, and of heat flow, are represented separately. At any one time, attention can be focused on the overall concept (that heat flow is caused by temperature difference), or on one or other detail (on either temperature difference or on heat-

flow). A related scheme for time-sharing in connectionist networks has been discussed by Hinton (1990).

However autonomous development of strategies requires a concept of the task, and this requires that there be sufficient processing capacity to represent the dimensions of the concept. Where children cannot represent sufficient dimensions for a particular concept, they will default to lower dimensionality representations, which will result in strategies that are partly correct, but which lead to errors on some variants of the task.

Model of serial processing strategy

In order to explore this aspect of cognitive development, we have produced a self-modifying production system model of transitive inference strategies (Halford et al., 1992, on contract). According to this model, a child uses a schema induced from ordinary life experience to provide a template for an ordered set. This template guides the development of strategies. However, as Figure 1 shows, the mapping of the problem into an ordering schema requires two relations to be processed, otherwise correspondence cannot be established.

If children cannot construct 3 dimensional representations, two relations cannot be processed, and strategies that are only partially valid will result. This typically results in errors such as the following: When the premise a>b is presented, the order ab is constructed. When b>c is presented, c is appended to ab, yielding the order abc. This is fine, but when a>c is presented, c is placed next to

a, yielding acb. A number of errors of this type arise from processing only one premise at a time, without integrating relations. This is typical of the performance of young children, and of older children and adults under high processing load (Maybery et al., in preparation).

Thus strategies reduce processing load, increase efficiency, and are an important component of cognitive development, as well as of expertise. They are not however panaceas, because development of strategies, unless taught exclusively through external input, requires that the child be able to represent the structure of the concept adequately.

Individual differences

This model of cognitive development provides three bases for individual differences. These are experience, processing capacity, and the interaction of the two. Because cognitive development is experience driven, and depends on accumulation of knowledge about the world, and acquisition of strategies and procedural knowledge, differences in opportunity for learning will inevitably affect development. The social environment clearly plays a major role in providing this experience, but social influences operate through the learning mechanisms that are built into the child. Chunking and segmentation are major acquisitions with experience, so these will depend on an individual's environment.

Individual differences in processing capacity probably operate through the clarity and effectiveness of representations. Less

ambiguous representations lead to faster solution times, less errors and, because they permit tensor products of higher rank to be computed without confusion, to representation of more complex relations. Vectors with a lot of "noise" or randomness will yield increasing ambiguity with higher rank tensor products, because of the complexity of interconnections involved.

The interaction of capacity and experience occurs because development of skills and strategies, as well as recoding processes such as chunking, depend on ability to represent the structure of concepts. If representations are of less dimensionality than required, this leads to strategies that are not effective in all circumstances. The result is partial competence, rather than genuine competence. The relation between learning and capacity has been discussed elsewhere (Halford, 1989b).

Cognitive growth

Cognitive growth depends therefore on four main factors:

The first is *learning and induction*, which enables the child to build up an extremely rich store of world knowledge. This is the "raw material" of the schemas which can be used as mental models in reasoning and problem solving.

The second factor is *conceptual chunking*, which entails recoding representations into fewer vectors, so they can be combined into more complex representations, without overloading processing capacity.

The third factor is the development of *serial processing strategies* which permit tasks to be performed in smaller steps, timesharing the available representational capacity.

The fourth factor is the *development of ability to represent concepts of higher dimensionality*. The first three factors are essentially experiential, but the fourth is probably at least partly maturational. The actual mechanism is not yet known, but it probably entails differentiating distributed representations into more vectors. This entails rearranging the connections, to make the representations equivalent to higher rank tensor products. It would not increase overall processing capacity, but would enable higher orders of relationship to be represented.

The type of change that is envisaged here is analogous to splitting an experimental design into more independent variables. The total number of conditions represented might not change, but the orders of interaction that can occur do change; e.g. if we take a two-way ANOVA with four levels of one factor and two levels of another, and convert it into a three factor design with two levels on each factor, we still have the same number of conditions (8), but now we have a three-way interaction as well as two-way interactions and main effects. Thus the most important change is in the orders of relations that can be represented. Similarly, growth in processing capacity through development is more likely to mean that higher order relations can be represented, rather than that more information can be stored.

The performance of a child who cannot construct representations of adequate dimensionality is analogous to analysing (say) a three-factor experiment as a series of two-way ANOVAS. Most findings will be a correct account of the data, just as the hypothetical child's performance will be mostly correct. There will be however, at least in certain telltale cases, higher order interactions that will be missed. Similarly, the child who deals with an N-dimensional concept using representations of dimensionality less than N is really looking at the task through restricted windows. Sooner or later telltale performances will occur which show that the representation was not really adequate.

We suggest that this is a good analog of the role of processing capacity in cognitive development. The differentiation of representations into more vectors, and therefore more dimensions, is an enabling factor that occurs at least partly through maturation, and which in turn enables children to construct strategies based on more adequate concepts of tasks. Thus cognitive development is an interaction of maturation which leads, inter alia, to representations of higher dimensionality, and experience which contributes to a knowledge base that provides mental models, schemas, and strategies, as well as restructured or chunked concepts that reduce the processing demands of tasks.

References

Anderson, J.R. (1987). Skill acquisition: compilation of weak-method problem solutions. Psychological Review, 94,192-210.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory. (Vol.2, pp. 89-195). New York: Academic Press.

Baddeley, A. (1990). Human memory: Theory and practice. Needham Heights, MA: Allyn and Bacon.

Best, J. B. (1992). Cognitive Psychology (3rd ed.) New York: West Publishing.

Carey, S. (1985). Conceptual change in childhood. Cambridge, MA: MIT Press.

Chi, M.H.T., & Ceci, S.J. (1987). Content knowledge: its role, representation and restructuring in memory development. Advances in Child Development and Behaviour, 20, 91-142.

Diamond, A. (1989). Development as progressive inhibitory control of action: retrieval of a contiguous object. Cognitive Development, 4, 223-249.

Foos, P.W., Smith, K.H., Sabol, M.A., & Mynatt, B.T. (1976). Constructive processes in simple linear order problems. Journal of Experimental psychology: Human learning and memory, 2, 759-766.

Goldman-Rakic, P.S. (1987). Development of cortical circuitry and cognitive function. Child Development, 58, 601-622.

Grossberg, S. (1980). How does brain build a cognitive code? Psychological Review, 87, 1-51.

Halford, G.S. (1982). The development of thought. Hillsdale, N.J.: Erlbaum.

Hlalford, G.S. (1989) Reflections on 25 years of Piagetian cognitive developmental psychology, 1963-1988. Human Development, 32, 325-387.

Halford, G.S. (1989b) Cognitive processing capacity and learning ability: An integration of two areas. Learning and individual differences, 1, 125-153.

Halford, G.S. (in press). Children's understanding: The development of mental models. Hillsdale, N.J.: Erlbaum.

Halford, G.S. & Leitch, E. (1989). Processing load constraints: a structure-mapping approach. In M.A. Luszcz & T. Nettelbeck (Eds.), Psychological development: perspectives across the life-span (pp. 151-159). Amsterdam: Elsevier.

Halford, G.S., Maybery, M.T., & Bain, J.D. (1986). Capacity limitations in children's reasoning: A dual task approach. Child Development, 57, 616-627.

Halford, G.S., Maybery, M.T., & Bain, J.D. (1988). Set-size effects in primary memory: An age-related capacity limitation? Memory and Cognition, 16, 480-487.

Halford, G.S., Smith, S.B., Maybery, M.T., Stewart, J.E.M., & Dickson, J.C. (1992) A computer simulation model of acquisition of transitive inference. Technical Report 92/1, Centre for Human Information Processing and Problem Solving, University of Queensland, Australia.

Halford, G.S., Smith, S.B., Dickson, J.C., Maybery, M.T., Kelly, M. E., Bain, J. D. & Stewart, J.E.M. (1992, on contract). Modelling the development of reasoning strategies: The role of analogy, knowledge and capacity. In Developing Cognitive Competence: New Approaches to Process Modelling. Hillsdale, NJ: Erlbaum.

Halford, G.S. & Wilson, W.H. (1980) A category theory approach to cognitive development. Cognitive Psychology, 12, 356-411.

Halford, G.S., Wilson, W.H., Guo, J., Gayler, R., Wiles, J., & Stewart, J.E.M. (in press). Connectionist implications for processing capacity limitations in analogies. In K.J. Holyoak & J. Barnden (Eds.), Advances in connectionist and neural computation theory, Vol. 2: Analogical Connections. Norwood, N.J.: Ablex.

Hinton, G.E. (1990). Mapping part-whole hierarchies into connectionist networks. Artificial intelligence, 46, 47-75.

Holland, J.H., Holyoak, K.J., Nisbett, R.E., & Thagard, P.R. (1986). Induction: Processes of inference, learning and discovery. Cambridge, MA: MIT Press.

Hunt, E., & Lansman, M. (1982) Individual differences in attention. In R.J. Sternberg (Ed.) Advances in the psychology of human intelligence. (vol. 1., pp. 207-254). Hillsdale, N.J.: Erlbaum.

Kail, R. (1991). Processing time declines exponentially during childhood and adolescence. Developmental Psychology, 27, 259-266.

Maybery, M.T., Bain, J.D., & Halford, G.S. (1986).
Information processing demands of transitive inference. Journal of Experimental Psychology: Learning, Memory and Cognition, 12, 600-613.

Maybery, M.T., Halford, G.S., Bain, J.D., & Kelly, M.E. (1988). Children's construction of series. Unpublished manuscript, University of Queensland, Brisbane Australia.

Miller, G.A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological review, 63, 81-97.

Palmer, S.E. (1978). Fundamental aspects of cognitive representation. In Rosch, E. & Lloyd, B.B. (Eds.), Cognition and categorization (pp. 259-303). Hillsdale, NJ: Erlbaum.

Piaget, J. (1950). The psychology of intelligence. (M. Piercy & D.E. Berlyne, Trans.) London: Routledge & Kegan Paul, (Original work published 1947).

Rescorla, R.A. (1988). Pavlovian conditioning: it's not what you think it is. American Psychologist, 43, 151-160.

Rudy, J.W. (in press). Elemental and configural associations, the hippocampus and development.
Developmental Psychobiology.

Rumelhart, D.E., & McClelland, J.L. (Eds.) (1986). Parallel distibuted processing: Explorations in the microstructure of cognition. Vol. 1: Foundations. Cambridge, MA: MIT Press.

Sheppard, J.L. (1978). A structural analysis of concrete operations. In J.A. Keats, K.F. Collis, & G.S. Halford, (Eds.) Cognitive

Development: Research based on a neo-Piagetian approach. London: Wiley.

Schneider, W., & Detweiler, M. (1987). A connectionist/control architecture for working memory. The psychology of learning and motivation, 21, 53-119.

Sternberg, R.J. (1980). The development of linear syllogistic reasoning. Journal of Experimental child psychology, 29, 340-356.

Thatcher, R.W., Walker, R.A., & Giudice, S. (1987). Human cerebral hemispheres develop at different rates and ages. Science, 236, 1110-1113.

Thayer, E.S. & Collyer, C.E. (1978). The development of transitive inference: A review of recent approaches. Psychological Bulletin, 85, 1327-1343.

Trabasso, T. (1977). The role of memory as a system in making transitive inferences. In R.V. Kail, Jr. & J.W. Hagen, Perspectives on the development of memory and cognition, (pp. 333-356). Hillsdale, NJ: Erlbaum.

Figure Captions

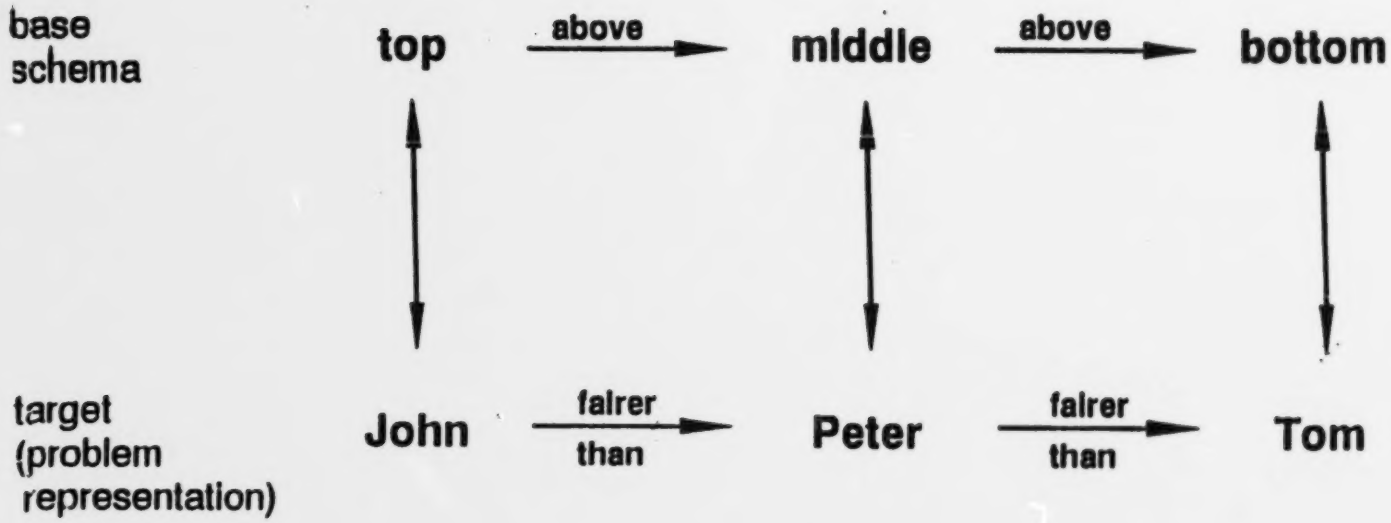
Figure 1. Transitive inference problem mapped into an ordering schema.

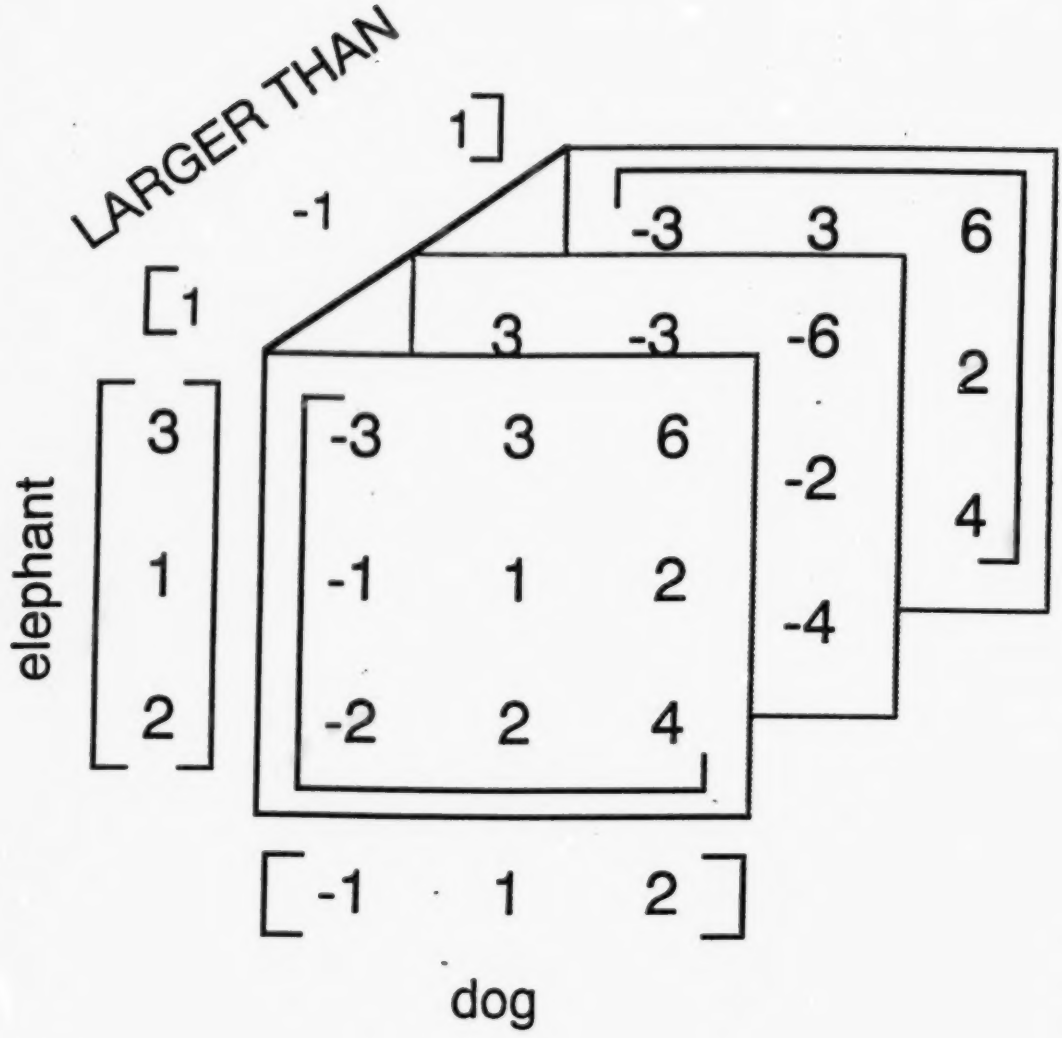
Figure 2. Tensor product representation of predicate-argument binding.

Figure 3. Tensor product representation of transitivity.

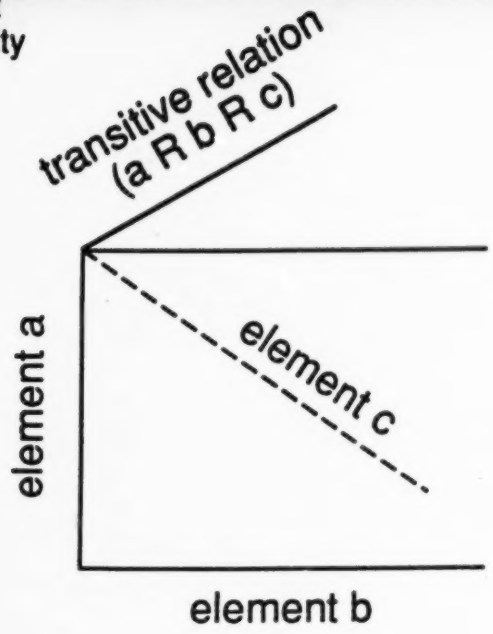
Figure 4. Dimensionality of representations related to tensor product representation and to Piagetian stage.

Premises: Peter is fairer than Tom
John is fairer than Peter

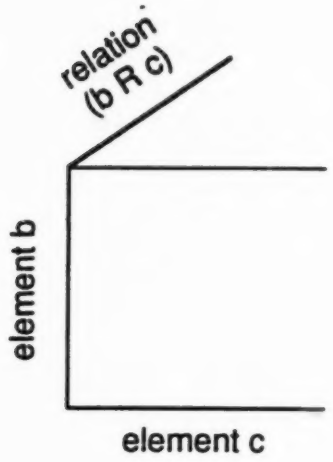


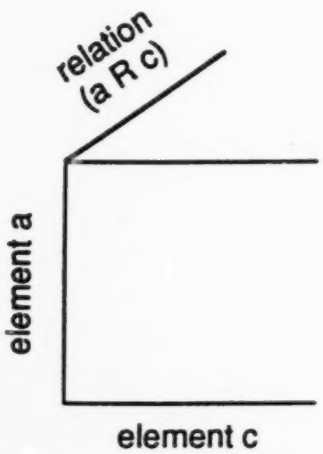

LARGER THAN
1
-1
1
3
-3
-6
-3
3
6
2
4
-2
-4
elephant
3
1
2
-3
3
6
-1
1
2
-2
2
4
-1
1
2
dog

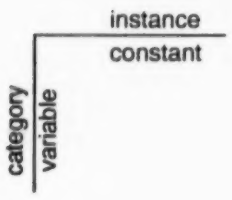
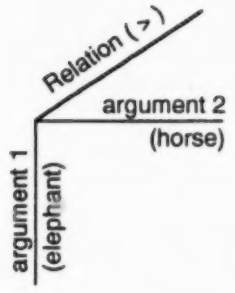
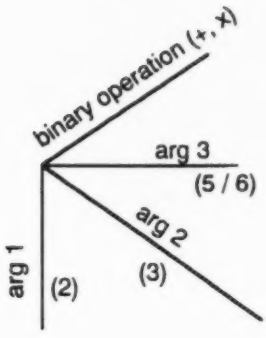
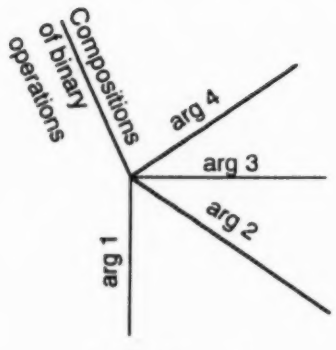
rank-4 tensor product
representing transitivity



rank-3 tensor products
representing premises





Piagetian Stage	Formal Specification	Dimensionality	PDP Implementation
Preconceptual	Unary relations	1	
Intuitive	Binary relations, univariate functions.	2	
Concrete Operational	Ternary relations, binary operations, bivariate functions.	3	
Formal Operations	Quaternary relations, compositions of binary operations.	4	

2 (2 + 3) = 10